SYSTEM TO PROVIDE CONSUMER PREFERENCE INFORMATION

BACKGROUND OF THE INVENTION

5

Field Of The Invention

The present invention relates to systems for determining consumer preferences. More specifically, the invention relates to systems for determining consumer preference information relating to product attributes and to product

10 attribute levels.

Description Of The Related Art

During the design of a product, a manufacturer must choose from among several available product features, or attributes, to include in the product. Some

15 attributes may be optional while others may be required. In the case of a television set, "Chassis color" is an attribute that must be included and "Picture-in-picture" is an optional attribute. For each included attribute, a manufacturer must also choose an attribute level to associate with the attribute. Attribute levels that may be associated with the attribute "Chassis color" include "black",

20 "white", "blue", etc.

Trade-off analysis techniques attempt to determine consumers' preferences for particular product attributes and attribute levels in order to identify ideal product configurations. A consumer, in this regard, is any entity to which a product may be offered. Such consumers include individuals,

25 businesses, and purchasing managers, and a product may include a good and/or service.

For example, trade-off analysis techniques allow a manufacturer to compare the attractiveness of a Sony television priced at $599 with that of a Magnavox television priced at $399. Such a comparison is possible because the

30 techniques associate a particular numerical value with a consumer's preference

for each attribute and attribute level. Accordingly, the relative attractiveness of differences or changes in attributes with respect to differences or changes in any other attribute can be determined simply by comparing the appropriate associated numerical values. For example, the attractiveness of a price change

5   from $599 to $399 may be compared with the attractiveness of a brand change from Magnavox to Sony. Therefore, by using consumer preference information, a manufacturer is more likely to choose product configurations as well as production amounts and prices for each product configuration that improve sales objectives such as overall profit, consumer satisfaction and consumer loyalty.

10   As described in the Background of commonly-assigned co-pending U.S. Patent Application Serial No. 09/754,612, entitled SYSTEM TO QUANTIFY CONSUMER PREFERENCES, which is incorporated by reference herein for all purposes, conventional trade-off analysis techniques include conjoint, discrete choice, self-explicated, and hybrid techniques. Each of these techniques may be

15   used to produce consumer preference information. However, these techniques often fail to produce a full complement of consumer preference information associated with a particular consumer. In other instances, the produced consumer preference information unsatisfactorily reflects the particular consumer's preferences. As a result, it is difficult to use conventionally-collected

20   consumer preference information to accurately determine, for example, an amount of change in a consumer's preference for a product that would result from a change in a particular attribute or a particular attribute level of the product.

In an attempt to address the foregoing, some conventional systems apply stabilization algorithms to the produced consumer preference information. The

25   stabilization algorithms are intended to improve the predictive precision and completeness of the consumer preference information. In one conventional system, the Adaptive Conjoint Analysis/Hierarchical Bayes module sold by Sawtooth Software, Inc., consumer preference information of other consumers is used to stabilize consumer preference information of a subject consumer.

30   However, these conventional stabilization algorithms are also not seen to produce sufficiently predictive or complete consumer preference information.

2

In view of the foregoing, what is needed is a system to determine consumer preference information that provides greater predictive precision than that produced by conventional systems.

5   SUMMARY OF THE INVENTION

In order to address the foregoing need, the present invention provides a system to determine consumer preference information in which preference information associated with a consumer is determined, and the preference information is mixed with preference information associated with a plurality of

10   consumers. According to this aspect, a degree to which the preference information associated with the consumer is mixed with the preference information associated with the plurality of consumers is different than a second degree to which second preference information associated with a second consumer is mixed with preference information of a second plurality of

15   consumers.

By virtue of the above features, the present invention may mix consumer preference information associated with a consumer with preference information associated with a group of consumers according to different degrees depending upon one or more factors. As a result, the present invention provides consumer

20   preference information which is more predictive and complete than that produced by previous systems. For example, in order to produce consumer preference information for a first consumer, consumer preference information associated with the first consumer may be mixed with consumer preference information associated with a first group of consumers in a 25/75 ratio. On the other hand, in

25   order to produce consumer preference information for a second consumer, consumer preference information associated with the second consumer may be mixed with consumer preference information associated with a second group of consumers in a 60/40 ratio. It should be noted that the first and second group of consumers may include all, some or no identical consumers. It should also be

30   noted that the mix may be represented by measures of degree other than the simple ratios of the previous example.

Further to the foregoing aspect, trade-off questions are provided to the consumer, actual answers to the trade-off questions are received, consumer answers to the trade-off questions are predicted based on the preference information associated with the consumer, and subgroup answers to the trade-off

5   questions are predicted based on the preference information associated with the plurality of consumers. Moreover, the preference information associated with the consumer is mixed with the preference information associated with the plurality of consumers based on the actual answers, the predicted consumer answers and the predicted subgroup answers. The steps of this further aspect allow additional

10   control and accuracy in determining the degree to which consumer preference information associated with the consumer should be mixed with consumer preference information associated with the plurality of consumers.

According to another aspect, the present invention relates to a system to determine preference information in which preference information associated

15   with a consumer is determined, the preference information is validated, and the preference information is mixed with preference information associated with a plurality of consumers based on the validating step. This aspect advantageously provides mixing that may differ among consumers. That is, because mixing according to this aspect is based on the validating step, mixing may differ in

20   cases where the validating step differs. As such, more appropriate mixing may be achieved than that achieved by previous systems.

In a related aspect, the validating step includes provision of trade-off questions to the consumer, reception of the consumer's actual answers to the trade-off questions, prediction of the consumer's answers to the trade-off

25   questions based on the preference information associated with the consumer and prediction of subgroup answers to the trade-off questions based on the preference information associated with the plurality of consumers. Moreover, the preference information is mixed with the preference information associated with the plurality of consumers based on the actual answers, the predicted consumer

30   answers and the predicted subgroup answers. This aspect provides even more

4

appropriate mixing that produces predictive and useful consumer preference information associated with a particular consumer.

According to yet another aspect, the present invention concerns a system to determine trade-off questions based on a plurality of attribute levels, each of

5   the plurality of attribute levels being associated with an attribute and a part worth utility value. In the system, attribute levels are grouped into objects including two attribute levels, each of the two attribute levels of an object being associated with different attributes, and the objects are grouped into pairs, each of which include two objects, a first object of a pair including a first two attribute levels associated

10   with two attributes and a second object of the pair including a second two attribute levels associated with the two attributes. A plurality of pairs on which to base the plurality of trade-off questions are selected from the pairs, wherein a first pair is more likely to be selected than a second pair if a sum of part worth utility values associated with each of the attribute levels of the first pair is greater

15   than a sum of part worth utility values associated with each of the attribute levels of the second pair. By virtue of this aspect, questions may be determined that deal with trade-offs in which a responding consumer is most interested, and allowing for thorough testing of the predictive precision of collected preference information.

20   With these and other advantages and features that will become hereafter apparent, a more complete understanding of the nature of the invention can be obtained by referring to the following detailed description and to the drawings appended hereto.

25   BRIEF DESCRIPTION OF THE DRAWINGS

FIG. 1 is a flow diagram of process steps to provide consumer preference information according to embodiments of the present invention.

FIG. 2 is a representative view of a tabular portion of a preference information database according to embodiments of the present invention.

30   FIG. 3 is a topographic view of a network architecture according to embodiments of the present invention.

FIG. 4 is a block diagram of an internal architecture of a central system according to embodiments of the present invention.

FIG. 5 is a flow diagram of process steps to validate consumer preference information according to embodiments of the present invention.

5　　　FIG. 6 is a representative view of a tabular portion of a preference information database according to embodiments of the present invention.

FIG. 7 is a representative view of a matrix used to determine trade-off questions according to one embodiment of the invention.

FIG. 8 is a representative view of a tabular portion of a validation answer

10　　database according to embodiments of the present invention.

FIG. 9 is a view of an interface used to present trade-off questions to a consumer according to embodiments of the present invention.

FIG. 10 is a flow diagram of process steps to determine preference information associated with a plurality of consumers according to embodiments

15　　of the present invention.

FIG. 11 is a representative view of a tabular portion of a preference information database according to embodiments of the present invention.

FIG. 12 is a representative view of a tabular portion of a preference information database according to embodiments of the present invention.

20　　FIG. 13 is a flow diagram of process steps to mix preference information associated with a consumer with preference information associated with a plurality of consumers according to embodiments of the present invention.

FIG. 14 is a representative view of a tabular portion of a preference information database according to embodiments of the present invention.

25　　FIG. 15 is a view of consumer preference information as presented to a client according to embodiments of the present invention.

DETAILED DESCRIPTION

FIG. 1 is a flow diagram of process steps 10 according to embodiments of

30　　the invention. Process steps 10 will be described briefly below in the interest of providing an immediate introduction to features of the present invention.

Accordingly, process steps 10 will be described later with respect to more specific examples and specific hardware and software embodiments, along with details of alternative embodiments.

Process steps 10 begin at step S1, in which preference information

5   associated with a consumer is determined. The preference information may be determined by retrieving stored preference information or by using a system such as that described in aforementioned U.S. Patent Application Serial No. 09/754,612 to collect consumer preference information. Of course, other systems to collect preference information may be used in step S1 in order to

10   determine preference information, with varying degrees of output quality resulting therefrom.

The preference information determined in step S1 may include information such as that shown in FIG. 2. In this regard, FIG. 2 is a view of a tabular representation of a portion of preference information database 100 according to

15   embodiments of the invention. As shown, the tabular portion includes several fields and several records associated with one or more of the fields. Identification field 110 indicates the consumer and the product associated with the tabular portion and also indicates a type of preference information stored in the tabular portion. In the present example, identification field 110 indicates that

20   the tabular portion stores "Raw" preference information. Different types of preference information that may be stored in preference information database 100 will be discussed in detail below.

The records in the tabular portion of FIG. 2 include attribute field 112 and attribute level/associated part worth value field 114. Attribute field 112 specifies

25   an attribute of the product specified in identification field 110, and attribute level/associated part worth value field 114 specifies attribute levels corresponding to an associated attribute as well as part worth values associated with each specified attribute level. A part worth value, as described in detail in Application Serial No. 09/754,612, is a value that represents a consumer's

30   preference, or utility, for an associated attribute level. A part worth value may therefore be used in order to compare a consumer's preference for one attribute

7

level of an attribute over a second attribute level of the attribute with the consumer's preference for a first attribute level of a second attribute over a second attribute level of the second attribute. More particularly, to the consumer reflected in FIG. 2, the attractiveness of the color yellow over the color green (6 −

5   0 = 6) is greater than the attractiveness of the brand K2 over the brand Fischer (5 − 2 = 3).

Some part worth values associated with attribute levels in preference information database 100 comprise the symbol "X". This symbol indicates that the associated attribute level is unacceptable to the associated consumer.

10   Stated differently, any product including the associated attribute level would be unacceptable to the consumer. In the present example, the associated consumer would be unwilling to purchase any Junior type downhill ski.

The attributes and attribute levels associated with a product in preference information database 100 may be determined based on information obtained

15   from a manufacturer of the product. Generally, the attributes and attribute levels are features for which the manufacturer wishes to obtain consumer preference information. It should be understood that, although the present disclosure primarily discusses manufacturers, the present invention may be utilized by sellers, distributors, market researchers or other parties interested in obtaining

20   consumer preference information.

It should be noted that the information stored in preference information database 100 for a particular product may reflect fewer or more attributes and/or attribute levels than shown in FIG. 2. Furthermore, it is contemplated that preference information database 100 may store data corresponding to multiple

25   consumers and to multiple products for each consumer. On the other hand, it is contemplated that preference information database 100 may store preference information associated with multiple consumers but corresponding only to those products to be sold by a particular manufacturer.

Preference information stored in preference information database 100

30   might not include a part worth value associated with each attribute level. Also, the preference information might not include unacceptable attribute levels. In

8

other words, the particular representation of preference information that is shown in FIG. 2 does not reflect all possible types and representations of preference information.

5   After the preference information is determined in step S1, the preference information is validated in step S2. Validation is generally a test to determine the predictive precision of the preference information with respect to the associated consumer. In one embodiment, validation of preference information includes presenting questions to the consumer, receiving actual answers to the questions, and predicting answers based on the preference information. In another
10   embodiment, validation of preference information further includes determining an extent to which the consumer's actual answers to the questions match the consumer's predicted answers.

Next, in step S3, the preference information is mixed with preference information associated with a plurality of consumers based on the validation. In
15   this regard, preference information database 100 may store average preference information representing a subgroup of past consumers. In one embodiment, the average preference information is mixed with the preference information determined in step S1 based on the actual answers to the questions presented during validation, the answers predicted based on the determined preference
20   information, and answers predicted based on the average preference information.

In some embodiments of step S3, the preference information associated with the consumer is mixed with the preference information associated with the plurality of consumers to a degree that differs from consumer to consumer. In a
25   specific example, preference information associated with a first consumer may be mixed with preference information associated with a plurality of consumers in a 60/40 ratio while preference information associated with a second consumer may be mixed with preference information associated with the same or a different plurality of consumers in a 20/80 ratio.

30   The mixed preference information resulting from step S3 may more accurately reflect preferences of the consumer than either the preference

information determined in step S1 or the preference information associated with the plurality of consumers. Consequently, process steps 10 may be used to produce consumer preference information which is more predictive and complete than that produced by previous systems.

5

Network Architecture

FIG. 3 is a topographic view of a network architecture according to embodiments of the present invention. Of course, many other architectures may be used to implement the invention. Shown in FIG. 3 is central system 200,

10 depicted as a mainframe computer. Central system 200 may be used to perform, for example, process steps 10 in order to determine preference information consisting of part worth values associated with a consumer and a product. Central system 200 may be operated by a company, such as assignee Blue Flame Data, Inc., providing trade-off analysis services to manufacturers and/or

15 other clients desiring to obtain consumer preference information.

In operation, central system 200 may use data input by consumers and clients, as well as legacy data, third party data and/or observed behavior data to produce consumer preference information. It should be noted that many other types of computing hardware may be used to perform the functions of central

20 system 200 described herein, including, but not limited to, a server, a workstation, a network, or any combination of one or more of the foregoing. Further details of central system 200 are set forth below with respect to FIG. 4.

In communication with central system 200 are several client devices 300. Client devices 300 according to the present invention may be operated by a

25 product manufacturer to transmit attributes and attribute levels for a given product to central system 200. In response, central system 200 may determine consumer preference information consisting of part worth values associated with each attribute and attribute level. Of course, central system 200 may determine attributes and attribute levels for a given product using data from other sources.

30 Client devices 300 may also receive information from central system 200 intended for display to a manufacturer or to another client. Such information may

include real-time monitoring of consumer answers, scenario simulations, and/or an interface allowing the operator to tweak existing thresholds or parameters while information is being gathered from consumers. Of course, the manufacturer may also use client device 300 to view consumer preference

5   information produced by and received from central system 200.

As shown in FIG. 3, client device 300 may include a server and/or a kiosk. Any other suitable device may be used as client device 300 according to the invention, including but not limited to a workstation, a mainframe computer, and a computer terminal. In the case that client device 300 is a device having its own

10   input and/or output devices, such as a kiosk, a consumer may also use client device 300 to input answers to questions posed in accordance with the invention and to input other indications to central system 200. Accordingly, client device 300 may be used to present an interface to the consumer that allows the consumer to input such information.

15   Information such as questions and answers may also be transmitted to and/or received from consumers as described above through consumer devices 400. Shown in FIG. 3 are consumer devices 400 represented by a telephone, a personal digital assistant, a workstation, and a pen-based computer. The illustrated connections indicate that the shown consumer devices 400 may

20   communicate with client devices 300, with client devices 300 and central system 200 and, in the case of telephone consumer device 400 or workstation consumer device 400, directly with central system 200. In this regard, consumer devices 400 usable in conjunction with the present invention include any device capable of presenting information to a consumer, visually and/or aurally, and of

25   transmitting an indication made by the consumer to an external device. Of course, consumer devices 400 should be able to communicate with the device or devices with which they are in communication over whatever type of network media exist between the devices.

Although the connections illustrated between the components of FIG. 3

30   appear dedicated, it should be noted that each of the connections may be shared by other components. Moreover, the connections may comprise one or more of a

local area network, a wide area network, a telephone network, a cellular network, a fiber-optic network, a satellite network, an infra-red network, a radio frequency network, or any other type of network which may be used to transmit information between devices. Additionally, the devices shown as in communication with

5 other devices need not be constantly exchanging data, rather, the communication may be established when necessary and severed at other times or always available but rarely used to transmit data.

Central System

10 FIG. 4 is a block diagram of the internal architecture of central system 200 according to embodiments of the invention. As illustrated, central system 200 includes microprocessor 210 in communication with communication bus 220. Microprocessor 210 may be a Pentium™, RISC™-based, or other type of processor and is used to execute processor-executable process steps so as to

15 control the components of central system 200 to provide desired functionality.

Also in communication with communication bus 220 is communication port 230. Communication port 230 is used to transmit data to and to receive data from external devices. Communication port 230 is therefore preferably configured with hardware suitable to physically interface with desired external

20 devices and/or network connections. In one embodiment, questions for consumers are transmitted to and answers are received from consumer devices 400 over communication port 230.

Input device 240, display 250 and printer 260 are also in communication with communication bus 220. Any known input device may be used as input

25 device 240, including a keyboard, mouse, touch pad, voice-recognition system, or any combination of these devices. Input device 240 may be used by an operator to input product-related information such as attributes and attribute levels, consumer-related information such as consumer preference information and contact information, client-related information such as billing and transaction

30 information, and commands to central system 200. In this regard, a command may be input to central system 200 to output a report detailing a particular client's

account, a particular consumer's preference information or preference information associated with a plurality of consumers.

Such a report may be output to display 250, which may be an integral or separate CRT display, flat-panel display or the like. Display 250 is used to output

5   graphics and text to an operator in response to commands issued by microprocessor 210. Printer 260 is also an output device, but produces a hardcopy of data using ink-jet, thermal, dot-matrix, laser, or other printing technologies.

RAM 270 is connected to communication bus 220 to provide

10 microprocessor 210 with fast data storage and retrieval. In this regard, processor-executable process steps being executed by microprocessor 210 are typically stored temporarily in RAM 270 and executed therefrom by microprocessor 210. ROM 280, in contrast, provides storage from which data can be retrieved but to which data cannot be stored. Accordingly, ROM 280 is

15 used to store invariant process steps and other data, such as basic input/output instructions and data used during system boot-up or to control communication port 230.

Data storage device 290 stores, among other data, central system program 292 of processor-executable process steps. According to embodiments

20 of the present invention, the process steps of central server program 292 may be read from a computer-readable medium, such as a floppy disk, a CD-ROM, a DVD-ROM, a Zip disk, a magnetic tape, or a signal encoding the process steps, and then stored in data storage device 290. Microprocessor 210 executes instructions of program 292 and thereby operates in accordance with the present

25 invention, and particularly in accordance with the process steps described in detail herein.

Specifically, according to embodiments of the invention, microprocessor 210 executes processor-executable process steps of central system program 292 to provide for determination of preference information associated with a

30 consumer, and mixing of the preference information with preference information associated with a plurality of consumers. The process steps of central system

13

program 292 are also executed according to these embodiments so that a degree to which the preference information associated with the consumer is mixed with the preference information associated with the plurality of consumers is different than a second degree to which second preference information

5   associated with a second consumer is mixed with preference information of a second plurality of consumers.

Also according to embodiments of the invention, the process steps are executed to determine preference information associated with a consumer, to validate the preference information, and to mix the preference information with

10   preference information associated with a plurality of consumers based on the validating step.

The foregoing aspects of the invention advantageously provide mixing that may differ among consumers. Since more appropriate mixing may be achieved than that achieved by previous systems, these aspects provide predictive and

15   useful consumer preference information.

Also included in central system program 292 may be processor-executable process steps to provide a World Wide Web server. Such a Web server would allow central server 200 to communicate with client devices 300 and consumer devices 400 through the World Wide Web. In addition, program

20   292 may include process steps of an interactive voice response system enabling central system 200 to transmit questions to and receive answers from a consumer using a telephone consumer device 400.

Central system program 292 may be stored in data storage device 290 in a compressed, uncompiled and/or encrypted format. In alternative embodiments,

25   hard-wired circuitry may be used in place of, or in combination with, processor-executable process steps for implementation of the processes of the present invention. Thus, embodiments of the present invention are not limited to any specific combination of hardware and software.

Also stored in data storage device 290 are preference information

30   database 100 and validation answer database 294. Preference information database 100 includes various types of preference information determined

according to the present invention. As will be described in more detail below, the types may include "raw" preference information, normalized preference information, and stabilized preference information associated with individual consumers, with subgroups of two or more consumers and/or with segments of

5   two or more consumers. Validation answer database 294 stores answers to trade-off questions presented according to the invention. The answers include actual answers of consumers, answers predicted based on preference information associated with consumers, and answers predicted based on preference information associated with a subgroup of two or more consumers.

10   Usage of preference information database 100 and validation answer database 294 is described in detail below. As will be understood by those skilled in the art, the tabular illustrations and accompanying descriptions of the databases merely represent relationships between stored information. A number of other arrangements may be employed besides those suggested by the tables

15   shown. Similarly, the illustrated entries of the databases represent sample information only; those skilled in the art will understand that the number and content of the entries can be different from those illustrated.

        Data storage device 290 also includes elements that may be necessary for operation of central system 200, such as other applications, data files, an

20   operating system, a database management system and "device drivers" for allowing microprocessor 210 to interface with devices in communication with communication port 230. These program elements are known to those skilled in the art, and are therefore not described in detail herein.

25   <u>Validation</u>

        FIG. 5 is a flow diagram of process steps 500 to validate preference information according to one embodiment of step S2 of process steps 10. Although process steps 10, process steps 500 and the other process steps described herein are described as being performed by central system 200

30   through execution of processor-executable process steps of central server program 292 by microprocessor 210, the process steps may also be performed,

in whole or in part, by one or more of central system 200, client devices 300, consumer devices 400, other devices, and manual means.

Process steps 500 begin at step S501, in which preference information associated with a consumer is determined. As mentioned above with respect to

5   step S1 of process steps 10, preference information may be determined in step S501 using the techniques described in U.S. Patent Application Serial No. 09/754,612 or using conventional techniques for determining preference information associated with a consumer. The preference information may also be determined in step S501 simply by receiving the preference information from

10 any source, such as client device 300, or by retrieving stored preference information associated with a consumer of interest. FIG. 6 illustrates a tabular representation of a portion of preference information database 100 storing preference information that may be determined in step S501 and that will be used to describe process steps 500.

15 After the preference information is determined in step S501, trade-off questions are determined based on the preference information in step S502. In some embodiments, the determined trade-off questions present difficult and relevant choices to the consumer based on the associated preference information. By presenting difficult and relevant choices to the consumer, the

20 predictive precision of the preference information may be better evaluated.

One example for determining trade-off questions is described below, but other systems may be used in accordance with the invention. According to the example, a combination of specific attribute levels of two attributes is referred to as an Object, and a combination of two Objects is referred to as a Pair. Based

25 on the FIG. 6 tabular portion, $550/All Mountain is a first Object, $450/Junior is a second Object, and a Pair may include the Objects $550/All Mountain and $450/Junior.

This embodiment attempts to create eight Pairs in which each Object of a Pair concerns a same two attributes, as in the above example in which both

30 Objects concern the attributes Price and Type. Moreover, for each of the eight Pairs, an attribute level of a first Object of a Pair is associated with a greater part

worth value than a corresponding attribute level (an attribute level associated with a same attribute as the attribute level of the first Object) of a second Object of the Pair, while a second attribute level of the first Object of the Pair is associated with a smaller part worth value than the other attribute level of the

5   second Object. The previously-described Pair, $550/All Mountain and $450/Junior, satisfies the foregoing guidelines, because the part worth value associated with All Mountain (1.82) is greater than the part worth value associated with Junior (1.04) and the part worth value associated with $550 (2) is less than the part worth value associated with $450 (4).

10   The present example uses additional guidelines based on which trade-off questions are determined in step S502. Specifically, a sum of part worth values associated with each attribute level of a first Object of a Pair should be as similar as possible to a sum of part worth values associated with each attribute level of a second Object of the Pair. Also, no one attribute should be represented in more

15   than three pairs, and no one Object should be present in more than two pairs. Each of these additional guidelines is intended to produce trade-off questions that thoroughly test the predictive precision of the preference information.

           According to another particularly inventive additional guideline, Applicants have discovered benefits resulting from identifying those Pairs in which the sum

20   of each part worth value of a Pair is largest and creating trade-off questions based on those Pairs. Such trade-off questions are believed to thoroughly test the predictive precision of the preference information by dealing with topics in which the responding consumer is most interested.

           One specific method for determining trade-off questions according to the

25   foregoing guidelines is hereafter described. According to the specific method, a matrix is created in which each row represents an attribute level of one attribute and each column represents an attribute level of a second attribute, with the attribute levels ordered from most to least preferred. Stored in each cell of the matrix is a sum of part worth values associated with the attribute levels

30   representing the row and column to which the cell belongs. Moreover, each cell

represents an Object consisting of the attribute levels representing the row and column to which the cell belongs.

Matrix 700 of FIG. 7 is an example of the above-described matrix created based on the data of FIG. 6. The sums associated with the attribute level $750

5   are shown as "N/A" in matrix 700 because $750 is an unacceptable attribute level. According to this specific method, a matrix such as matrix 700 is created for every possible combination of two attributes. In the present example, matrices are created for the combinations color/price (matrix 700), color/brand, color/type, price/brand, price/type, and brand/type. In other embodiments, matrix

10   700 is only created for combinations of highest-ranked attributes. Also, according to some embodiments, matrix 700 is populated only with highest-ranked attribute levels.

Each matrix is examined to identify Pairs of Objects for which an attribute level of a first Object is associated with a greater part worth value than a

15   corresponding attribute level of a second Object, while a second attribute level of the first Object is associated with a smaller part worth value than the other attribute level of the second Object. One method for such identification includes selecting a cell in the matrix. Next, each cell "northeast" of the selected cell is identified. Accordingly, the Object represented by the selected cell can be paired

20   with the object represented by any of the identified cells to create a Pair satisfying the foregoing criteria.

For all such Pairs identified from each matrix, the difference in the summed part worth values representing each Object of a Pair is calculated. For example, in the case of the Pair consisting of Objects Yellow/$450 and

25   Blue/$550, the calculated difference based on matrix 700 equals 4.8. Next, each Pair for which the calculated difference is greater than two is discarded. The remaining Pairs are then ranked according to the sums of all the part worth values associated with their Objects, with the Pair having the highest sum being ranked first. If the sums corresponding to two Pairs are equal, the Pair having

30   the smallest difference in the summed part worth values representing each Object is ranked above the other Pair.

Eight Pairs are then selected from the top of the ranked list. If one
attribute is represented more than three times among the eight Pairs, the lowest-
ranking Pair(s) of the selected Pairs which represent the attribute is ignored and
a next-ranked Pair is selected. If less than eight Pairs are selectable, the

5   selection is repeated with an attribute being allowed to be represented no more
than four times. If less than eight Pairs are again selected, selection re-occurs
with an Object being allowed to be chosen up to three times. If still less than
eight Pairs are chosen, all Pairs for which the calculated difference was three or
less are re-ranked and selection of the re-ranked Pairs proceeds as described

10   above.

       Once the eight Pairs are selected, trade-off questions are determined
therefrom. FIG. 8 illustrates a tabular representation of a portion of validation
answer database 294 including Pairs selected according to step S502. As
shown, each row represents a selected Pair and therefore also represents a

15   trade-off question. For example, trade-off question 1 requires a consumer to
compare his preference for a $550 All Mountain downhill ski with his preference
for a $450 Junior downhill ski. The trade-off questions are stored and presented
to a consumer in random order, with the Objects of each Pair also randomly
appearing as first or second Objects in a Pair.

20   Each trade-off question represented in validation answer database 294 is
associated with fields for specifying an actual preference intensity, a predicted
preference intensity, and a predicted subgroup preference intensity. Use of
these latter three fields is described in detail below.

       Many systems other than those described above may be used to

25   determine trade-off questions according to step S502. These systems may use
any number of the above principles and techniques.

       Returning to process steps 500, the trade-off questions determined in step
S502 are presented to a consumer in step S503. The consumer to whom the
questions are presented is preferably the consumer associated with the

30   preference information determined in step S501. The questions may be
presented by transmitting data to client device 300, consumer device 400, or

another device operated by the consumer, by creating a hardcopy questionnaire to be mailed to the consumer, or by using any other known means of presenting information. FIG. 9 illustrates one embodiment for presenting trade-off questions according to step S503. According to the illustrated embodiment, data

5   representing trade-off question 1 of validation answer database 294 is transmitted to consumer device 400 and displayed by a display device thereof.

After the trade-off questions have been presented to the consumer, actual answers to the trade-off questions are received in step S504. In this regard, the actual answers may be transmitted by consumer device 400 and received by

10   communication port 230 of central system 200, input by an operator operating input device 240 of central system 200, or received by other means. The actual answers according to the present example comprise currency-normalized preference intensities.

As shown in FIG. 9, the consumer is asked to indicate a preference

15   intensity for one Object over another Object. In response, the consumer operates consumer device 400 to indicate a preference intensity and to transmit the preference intensity to central system 200. Upon receipt of the preference intensity, the preference intensity may be currency-normalized by multiplying it with a conversion factor associated with the consumer's preference information,

20   and then stored in association with the appropriate trade-off question in the actual preference intensity field of validation answer database 294. Therefore, if the sliding bar of FIG. 9 is moved to the "-3" indicator and if the conversion factor equals $50/util, the actual preference intensity is −150. This scenario is reflected in the first illustrated record of validation answer database 294. Of course, this

25   process is repeated for each of the determined trade-off questions. Currency-normalization is discussed in more detail below.

In step S505, answers to the trade-off questions are predicted based on the preference information determined in step S501. In one embodiment of step S505, the answers are predicted by subtracting the sum of part worth values

30   associated with Object 1 of a trade-off question from the sum of part worth values associated with Object 2 of the trade-off question. Using trade-off question 1 of

FIG. 8 and the part worth values shown in FIG. 6 as an example, a predicted, non-currency-normalized preference intensity according to this embodiment is equal to (2 + 1.82) - (4 + 1.04) = -1.22. After step S505, the predicted preference intensities are currency-normalized and stored in association with appropriate

5   trade-off questions in the predicted preference intensity field of validation answer database 294.

Stabilization

According to the present invention, stabilization refers to the mixing of

10   preference information associated with a consumer with preference information associated with a plurality of consumers to produce mixed preference information. As described with respect to process steps 10, stabilization follows validation according to one embodiment of the invention. One example according to this embodiment is set forth in process steps 1000 of FIG. 10 and

15   process steps 1300 of FIG. 13.

In the particular embodiment described below, the plurality of consumers includes all past consumers for whom associated preference information is stored in preference information database 100. In other embodiments, the plurality of consumers includes a predetermined number of past consumers,

20   such as the last five thousand consumers for whom associated preference information was stored in preference information database 100. The plurality of consumers may also include all consumers for whom associated preference information was stored in preference information database 100 during a particular time period.

25   According to process steps 1000, all preference information associated with past consumers is currency-normalized in step S1001. Currency-normalization is performed because of the impossibility in comparing raw preference information associated with one consumer to raw preference information associated with another consumer or consumers. Comparison is

30   impossible because individual consumers have individual opinions on the weight of values in a scale. In other words, a first consumer may assign a preference

value of 8 to an attribute level, while a second consumer who equally prefers the attribute level may assign a preference value of 6. In order to allow these comparisons and to establish a consistent relationship between the preference information associated with each of the past consumers, the present inventors

5   have discovered that part worth values of each set of preference information associated with past consumers should be calibrated to a scale based on currency. This discovery takes advantage of the realization that relative preferences for different amounts of currency will not vary widely among a group of consumers. It should be noted that other systems for normalizing the

10   preference information associated with the past consumers may be used in step S1001, such as systems using a purchase likelihood scale.

In order to currency-normalize preference information according to the present embodiment, each set of preference information associated with past consumers is first obtained from preference information database 100. In this

15   regard, preference information database 100 stores preference information associated with individual past consumers in addition to the consumer for whom preference information was determined in step S501. Next, each part worth value in a given set of preference information associated with an individual past consumer is multiplied by the following conversion factor:

20

$$\frac{abs(Price\ max - Price\ min)}{abs(part\ worth\ value\ for\ Price\ max - part\ worth\ value\ for\ Price\ min)}.$$

For example, the conversion factor corresponding to the raw preference

25   information of FIG. 6 is equal to abs[($650 − $450)/(0 utils − 4 utils)] = $50/util. Each part worth value in the given set of preference information is multiplied by the conversion factor in order to produce currency-normalized preference information. FIG. 11 shows the preference information of FIG. 6 after currency-normalization as described above. The currency-normalized preference

30   information is stored in preference information database 100, with identification

22

field 110 specifying that the associated preference information has been currency-normalized.

It should be noted that different conversion factors may be used to currency-normalize different sets of preference information because the

5   conversion factor for a given set of preference information is based on part worth values included in the given set of preference information. Moreover, a conversion factor according to the present invention may comprise a constant or an equation such as a distribution, piecewise function, or the like.

After each set of preference information associated with the individual past

10   consumers is currency-normalized, first optimal segments are determined based on unacceptable attribute levels in step S1002. More specifically, the past consumers are grouped into segments based on attribute levels that the past consumers indicated as being acceptable and unacceptable. Grouping on this basis is intended to ensure that consumers in a segment are associated with

15   similar preference information with respect to acceptable and unacceptable attribute levels.

The first optimal segments may be determined in step S1002 using traditional cluster analysis (k-means method) or mixture models. If mixture models are used in step S1002, a binomial distribution may be used for segment

20   densities.

Detailed algorithm No. 1, set forth after the present Detailed Description, describes a system for performing stabilization according to one specific embodiment of the invention. Step 1 of the algorithm corresponds to step S1002 of process steps 1000. As will be understood after considering the algorithm,

25   step 1 describes a method in which latent cluster analysis is initially performed assuming two segments of past consumers. According to step 1, latent cluster analysis is then repeatedly performed assuming different numbers of segments.

After latent cluster analysis, a procedure known as CAIC scoring is used to determine which of the assumed number of segments is optimal. In this

30   regard, fit of the segments to past consumer preference information becomes more accurate as the number of segments increases, but the number of

parameters that must be estimated also increases. CAIC scoring considers both of these effects in determining the first optimal segments. Once the first optimal segments are determined, a probability that each past consumer belongs to each of the optimal segments is determined and stored.

5   In step S1003, second optimal segments are determined based on part worth values of the preference information that was currency-normalized in step S1001. Step S1003 is intended to group consumers having similar part worth values associated with acceptable attribute levels. Accordingly, in one embodiment of step S1003, unacceptable attribute levels are ignored. In another

10   embodiment, grouping is based on the belief that a consumer associated with a part worth value of 5 for an attribute level is more similar to a consumer for whom the attribute level is unacceptable than is a consumer associated with a part worth value of 10 for the attribute level.

As described with respect to step S1002, the determination of second

15   optimal segments may be performed using traditional k-means analysis or using mixture models. In some embodiments using mixture models, standard normal densities are assumed. If using a traditional analysis, ignoring the unacceptable attribute levels results in a "distance measure" between a consumer and a "cluster center" that is based on currency-normalized part worth values

20   associated with acceptable attribute levels. If using mixture models, a density value corresponding to the unacceptable attribute levels is ignored.

Step 2 of detailed algorithm no. 1 sets forth one specific system for performing step S1003 using latent cluster analysis. According to step 2, latent cluster analysis is performed by assigning, for each set of preference information

25   associated with past consumers, acceptable attribute levels to their associated part worth value and ignoring unacceptable attribute levels.

Step 2 continues similarly to step 1 of detailed algorithm no. 1, with latent cluster analysis being initially performed assuming two segments of past consumers and then repeatedly performed assuming different numbers of

30   segments. CAIC scoring is then used to determine which of the assumed number of segments is optimal based on results of the latent cluster analyses.

24

As a result of step 2, a probability that each past consumer belongs to each of the second optimal segments is determined and stored.

It should be noted that systems other than CAIC scoring may be used to determine an optimal number of segments in accordance with the present

5   invention. Detailed algorithm no. 2, set forth after detailed algorithm no. 1, demonstrates a system using for determining an optimal number of segments based on entropy and CAIC scoring.

Subgroups are determined based on the first and second optimal segments in step S1004. For example, it is assumed that four optimal segments

10 (A,B,C,D) are determined in step S1002 and three optimal segments (I,II,III) are determined in step S1003. Twelve possible subgroups (AI,AII,AIII,BI,BI,BIII,CI,CII,CIII,DI,DII,DIII) are identifiable based on these segments. One embodiment of step S1004 is set forth in step 3 of detailed algorithm no. 1. It should be noted that the number of subgroups may vary.

15 Detailed algorithm no. 1 presents an example including four subgroups.

In another embodiment of steps S1002 through S1004, first optimal segments are determined (e.g., A,B,C) and second optimal segments are determined based on the first optimal segments. According to one example of this embodiment, the subgroups determined in step S1004 are subgroups AI, AII,

20 BIII, BIV, BV, CVI, and CVII.

Next, in step S1005, average currency-normalized preference information is determined for each subgroup determined in step S1004. Step 4 of detailed algorithm no. 1 illustrates one embodiment of step S1005. Generally, currency-normalized part worth values of all past consumers in a subgroup are identified,

25 and those part worth values that are associated with a particular attribute level are averaged to determine an average currency-normalized part worth value for the particular attribute level. Then, part worth values that are associated with each other particular attribute level are averaged to determine an average currency-normalized part worth value for each other particular attribute level.

30 The process is repeated for each subgroup.

FIG. 12 illustrates a tabular portion of preference information database 100 storing average currency-normalized preference information for subgroup D2 determined according to step S1005. As shown, identification field 110 specifies that the information is associated with a particular subgroup and that the

5   preference information has been currency-normalized.

It should be noted that, according to some embodiments, process steps 1000 are performed periodically in order to maintain an up-to-date set of subgroups and associated average preference information. For example, subgroups may be re-determined every three weeks, whenever preference

10   information for one thousand new consumers is received, or according to some other criteria. In other embodiments, process steps 1000 are performed each time stabilized preference information associated with a consumer is desired. As a result of the re-determination of subgroups, two past consumers may each be associated with a same subgroup prior to a re-determination and associated with

15   respective different subgroups after the re-determination.

FIG. 13 is a flow diagram of process steps 1300. In some embodiments, process steps 1300 are performed each time stabilized preference information associated with a consumer is desired. As mentioned above, process steps 1300 may be performed separately from periodic performances of process steps

20   1000 or in conjunction with process steps 1000 as a single process to produce preference information associated with a particular consumer.

Flow begins at step S1301, in which preference information associated with a consumer is currency-normalized. Currency-normalization may proceed as described with respect to step S1001, and produces preference information

25   such as that illustrated in FIG. 11. Of course, FIG. 11 was described previously as illustrating currency-normalized preference information associated with a past consumer. In the present instance, it is assumed that the consumer represented in FIG. 11 is a "current" consumer, i.e., a consumer for which associated stabilized preference information is to be produced.

30   In step S1302, an optimal subgroup to which the consumer belongs is determined based on the preference information that was currency-normalized in

step S1301. Step 5.1 of detailed algorithm no. 1 describes a particular method for determining the optimal subgroup based on similarities between the consumer's unacceptable/acceptable attribute levels and on similarities between currency-normalized part worth values associated with the consumer and those

5   associated with the subgroup.

An optimal mixture of the preference information produced in step S1301 and the average preference information of the optimal subgroup determined in step S1005 is determined in step S1303. According to one embodiment, the determination of step S1303 is based on the actual answers to the trade-off

10   questions provided by the current consumer and the answers predicted based on the current consumer's preference information in step S505. Moreover, the determination is based on answers to the trade-off questions predicted based on the average currency-normalized preference information of the optimal subgroup. Accordingly, the latter answers are predicted based on the average currency-

15   normalized preference information of the optimal subgroup as described in step S505. As shown in FIG. 8, each of these types of answers may be stored in validation answer database 294 for convenient reference and retrieval. Step 5.2 of detailed algorithm no. 1 sets forth one system for determining an optimal mixture according to this embodiment.

20   Lastly, in step S1304, the currency-normalized preference information associated with the current consumer is mixed with the average currency-normalized preference information associated with the optimal subgroup according to the optimal mixture. In one embodiment of step S1304, each part worth value of the consumer's preference information is mixed with a

25   corresponding part worth value of the subgroup's preference information according to the optimal mixture. A method according to this embodiment is described in step 5.3 of detailed algorithm no. 1.

FIG. 14 is a tabular representation of a portion of preference information database 100. As specified by identification field 110, the portion includes

30   preference information associated with a particular consumer and stabilized in accordance with the present invention.

It may be determined in step S1303 that the optimal mixture is equal to 100% of the current consumer's preference information and 0% of the subgroup's preference information. In these instances, mixing the current consumer's preference information with the subgroup's preference information

5   does not produce more predictive preference information than that determined in step S501. More specifically, the currency-normalized preference information of the current consumer is identical to the stabilized preference structure of the current consumer.

FIG. 15 is a view of preference information 1500 as presented to a client

10   according to embodiments of the present invention. More particularly, preference information 1500 is a matrix of currency-normalized part worth values and associated attribute levels stabilized according to the present invention. Preference information 1500 is intended to provide a client with a comprehensible breakdown of preference information associated with a particular

15   consumer and determined according to the present invention. Preference information 1500 may be presented to a client in many ways, including by transmitting data representing preference information 1500 to client device 300, by transmitting a Web page including preference information 1500 to client device 300, by displaying preference information 1500 to the client using display

20   250, and by providing to the client a hardcopy of preference information 1500 produced using printer 260. As shown, preference information 1500 reflects the data stored in the portion of preference information database 100 represented in FIG. 14.

A client may determine an offer based on preference information 1500. In

25   a particularly advantageous embodiment, the offer is then provided to the consumer associated with preference information 1500. Such a personalized offer may be more likely to be accepted and may generate more profit than an offer designed for the general public. In some embodiments, a stabilized preference structure such as preference information 1500 is used to calculate

30   other useful information such as an optimal product selected from a product line.

Although the present invention has been described with respect to particular embodiments thereof, those skilled in the art will note that various substitutions may be made to those embodiments described herein without departing from the spirit and scope of the present invention.

Detailed algorithm no. 1

N = a total number of attribute levels that are either acceptable or unacceptable to a consumer.

n = a specific attribute level, $n = 1,...,N$

I = a total number of segments

i = a specific segment, $I = 1,...,I$

R = a total number of consumers in all segments

r = a number identifying a specific consumer, $r = 1,...,R$

S_i = a number of consumers in segment i.

s = a specific consumer of segment i, $s = 1,...,S_i$

$\rightarrow$ = vector

$\| \ \|$ = matrix

Σk = summation over the variable k

Πk = multiplication over the variable k

1.   Determine First Optimal Segments Based On Unacceptable Attribute Levels

1.1 Latent Cluster Analysis

$\| Y \|$ = a matrix that corresponds to a particular set of preference information associated with each consumer r.

$\| Y \| = \{Y1\rightarrow, Y2\rightarrow, ..., YR\rightarrow\}$

$Yr\rightarrow = (Y1,r)$

$\qquad (\quad . \quad)$

$\qquad (\quad . \quad)$

$$(\quad . \quad)$$
$$(Y_{N,r})$$

The following table illustrates a set of raw part worth values associated
with hypothetical past consumer r = 33.

Attribute 1	2	3	4
unacceptable level	7	10	5
0	4	4	2
6	3	0	0
no level	0	1	6

If a level n is unacceptable to the consumer, then $Y_{n,r} = 0$. If the level n is
acceptable to the consumer, $Y_{n,r} = 1$. Accordingly, the above set of part worth
values corresponds to $Y_{33\to} = \parallel 0\ 1\ 1\ 1\ 1\ 1\ 1\ 1\ 1\ 1\ 1\ 1\ 1\ 1\ 1\ 1 \parallel'$ (the single
quote denotes a conversion from row vector to column vector). Note that the last
row corresponding to Attribute 1 is ignored because Attribute 1 has only three
levels.

1.1.1   Perform latent cluster analysis assuming I = 2 segments

As an initial condition for the latent cluster analysis, randomly distribute
the past consumers among the 2 segments. For example, begin with the
following random initial conditions:

r = r1, r2, r15, ... are in segment i = 1 and r = r3, r4, r22, ... are in segment i = 2.

1.1.2 Calculate $\parallel \Theta \parallel$

$\parallel \Theta \parallel = (\theta_{1\to}, ..., \theta_{l\to})$

$\theta_{i\to} = (\theta_{i,1})$

$$(\quad . \quad)$$
$$(\quad . \quad)$$
$$(\quad . \quad)$$

(θi,N),

where θi,n = (Σr Yi,n,r * Ind i,r)/Si, where Ind i,r = an indicator function that equals 1 if a consumer r belongs to segment i, but which equals 0 if consumer r does not belong to segment i. Accordingly, the summation is performed only for
5   those consumers that belong to segment i.

The following table illustrates the calculation of θ1→, assuming that consumers r=1, r=2, etc. belong to segment 1.

Y1→	Y2→	Y15→	Y...→	θ1→ = average of Y values for consumers of segment 1
1	1	1	...	.12
1	1	1	...	.03
1	1	1	...	.25
1	1	1	...	1
1	1	1	...	.25
1	1	1	...	.12
0	1	1	...	0.5
1	1	1	...	0.5
1	1	0	...	.06
1	1	1	...	.03
1	1	1	...	.06
1	1	1	...	.06
0	1	1	...	0.5
1	1	1	...	.06
1	1	0	...	.03

Next, $\theta2\rightarrow$ is calculated using the Yn,r values of the members of segment 2.

1.1.3 Calculate $\|f\|$

5      $\|f\|$ = f1,1$\rightarrow$, f2,1$\rightarrow$, f1,2$\rightarrow$, f2,2$\rightarrow$,..., fl,R$\rightarrow$

There are (R)*(I), f i,r$\rightarrow$ vectors:

fi,r$\rightarrow$ = (fi,1,r)

...

(fi,N,r),

10   where fi,n,r = [(θi,n)^Yn,r]*(1-θi,n)^(1-Yn,r)

The following table illustrates an example of the calculation of fi,r$\rightarrow$.

f1,1$\rightarrow$ = f$\rightarrow$ for r=1 of S1   = (θ1$\rightarrow$ ^ Y1$\rightarrow$) *   ((1$\rightarrow$ - θ1$\rightarrow$)^(1$\rightarrow$ -Y1$\rightarrow$))
0.12
0.03
0.25
1.00
0.25
0.12
0.50
0.50
0.06
0.03
0.06
0.06
0.50
0.06
0.03

where $1\rightarrow$ is an identity vector.

Again assuming 2 segments and R consumers, also calculate $f2,1\rightarrow$,

5    $f1,2\rightarrow$, $f2,2\rightarrow$, $f1,3\rightarrow$, $f2,3\rightarrow$, $f1,4\rightarrow$, $f2,4\rightarrow$, $f1,...\rightarrow$, $f2,...\rightarrow$, $f1,R\rightarrow$, and $f2,R\rightarrow$.

1.1.4 Calculate $F\rightarrow$

$F\rightarrow = [F1,1,\ F2,1, F1,2, F2,2,...,FI,R]$

Fi,r is a value based upon fi,n,r. There are $(R)*(I)$ Fi,r values. Fi,r = the

10   product of the components of $fi,r\rightarrow$. $Fi,r= (f i,1,r) * (fi,2,r) * (fi,3,r) * ... * (f i,N,r)$.

For example, $(F1,1) = .12*.03*...*.03 = 6.56*10^{(-13)}$

Again assuming 2 segments and R consumers, also calculate F2,1, F1,2,

F2,2, F1,3, F2,3, F1,4, F2,4, F1,..., F2,..., F1,R, and F2,R.

15

1.1.5 Calculate Posterior Probability $P\rightarrow$

$P\rightarrow = [P1,1\ P2,1\ P1,2\ P2,2\ ...\ PI,R]$

Pi,r = the probability that consumer t belongs to segment i. There are

20   $(R)*(I)$ posterior probability values. $Pi,r = (Si*Fi,r)/ (\Sigma i\ Si*Fi,r)$

As an example, assume $(F2,1) = 7.11*10^{(-11)}$, I=2. Accordingly, $(P1,1) =$
$(S1*F1,1) / [(S1*F1,1) + (S2*F2,1)] = (S1*6.56*10^{(-13)}) /[S1*6.56*10^{(-13)} +$
$S2*7.11*10^{(-11)}] = 9.14*10^{(-3)}$.

25

Because there are only two segments and the probability that r = 1

belongs to either segment 1 or segment 2 equals 1, and $(P1,1) \approx 0$, then $(P2,1)$

≈ 1. Accordingly, the probability that r = 1 belongs to segment 1 is approximately

0 while the probability that C1 belongs to segment 2 is approximately 1. Next,

30   calculate the following assuming there are 2 segments and R consumers: (P1,2);

(P2,2); (P1,3); (P2,3); (P1,4); (P2,4); (P1,...); (P2,...); (P1, R); and (P2, R).

1.1.6 Calculate the likelihood L

There is one L value. $L = \Sigma i,r \, (\ln(Fi,r) * Pi,r)$

For example, assume:

5 $(F1,1) = 6.56*10^{\wedge}(-13)$

$(F2,1) = 7.11*10^{\wedge}(-11)$

$(P1,1) = 9.14*10^{\wedge}(-3)$

Therefore,

$L = \ln((6.56*10^{\wedge}(-13)))*(9.14*10^{\wedge}(-3)) + \ln((7.11*10^{\wedge}(-11)))*(...) + ...+ = .94.$

10

1.1.7 Iteration

Repeat steps 1.1.3 thru 1.1.6 in an iterative fashion until the posterior probabilities substantially converge, e.g., abs(Pi,r for iteration j – Pi,r for iteration j-1) < .005.

15

1.1.7.1 Calculate new $\|\Theta\| = (\theta1\rightarrow,...,\theta l\rightarrow)$

Instead of using the $\|\Theta\|$ formula of step 1.1.2 use the following formula for $\theta i\rightarrow$ for all repetitions of steps 1.1.3 through 1.1.6 until iteration ends.

20 $\theta1,n = [\Sigma r \ (P1,r * Yn,r)] / [\Sigma r \ (P1,r)]$

$\theta i\rightarrow = \{ \qquad \Sigma r \ (Pi,r * Y1,r) / [\Sigma r \ (Pi,r)]$

$$...$$

$\Sigma r \ (Pi,r * YN,r) / [\Sigma r \ (Pi,r)] \qquad \},$

where Pi,r and Yn,r are values from an immediately previous iteration.

25

For example, assume:

$(P1,1) = 9.14*10^{\wedge}(-3)$

$(Y1,1) = 1$

$(Y1,2) = 1$

30 $(Y1,...) = ...$

$(Y1,...) = ...$

...

Therefore,

$\theta1,1 = ((9.14*10^\wedge(-3))*1 + ...*1 + ...) / (9.14*10^\wedge(-3) + ...) = .23$.

Repeat for all N rows of the vector to obtain $\theta1\rightarrow$. Next, repeat to obtain $\theta2\rightarrow$.

5

1.1.8 Ensure global Maximum

Repeat steps 1.1.1 thru 1.1.7 using a new set of initial conditions. For example, try new random initial conditions specifying that r = 1, r = 3, r = 15, ... are in segment i =1 and r = 2, r = 4, r = 22, ... are in segment i = 2.

10

1.1.9 Repeat for other numbers of segments I

1.1.1 thru 1.1.8 were performed above for I = 2. Perform steps 1.1.1 thru 1.1.8 for I = 1, for I = 3, I = I, etc.

15 ## 1.2 CAIC scoring

1.2.1 Calculate a CAIC score for I = 2 segments

Identify the final L (likelihood) of each of the 2 iterations. Take the final L that is greater of the two. Also take the final posterior probabilities that correspond to that final, highest L. Using these values, CAIC = $(-2)*(\ln L) +$

20 $(N)*(I)* (\ln R)$.

1.2.2 Calculate a CAIC score for all other numbers of segments I

Identify the minimum CAIC score of the CAIC scores for all other numbers

25 of segments I, as well as I = 0. This score corresponds to the first optimal number of segments for the unacceptable latent cluster analysis. For example, if the CAIC scores corresponding to 2, 3 and 4 segments are 40.3, 60.6 and 80.4, respectively, then the first optimal number of segments is 2.

30

2. Determine Second Optimal Segments Based on Currency-Normalized Part Worth Values

2.1 Latent Cluster Analysis

5    $\|Y\|$ = a matrix that corresponds to a set of preference information for each consumer.

$\|Y\|$ = {Y1→, Y2→,..., YR→ }

Each level of Yr→ corresponds to the currency-normalized part worth value ($PW) of the associated attribute level for consumer r. However, if the

10   associated attribute level is unacceptable, the associated attribute level is ignored.

2.1.1 Perform latent cluster analysis assuming I = 2 segments

As an initial condition for the latent cluster analysis, randomly distribute

15   the consumers among the I = 2 segments. For example, begin with the random initial conditions that r = 1, r = 2, r = 15, ... are in segment i =1 and r = 3, r = 4, r = 22, ... are in segment i = 2.

2.1.2 Calculate $\|M\|$

20    $\|M\|$ = {Mi→}= { M1→ ,M2→ }, if I = 2

Mi,n = (Σt Yn,r * Ind)/Si

Mi→ = (Σt Yr→ * Ind)/Si

Y1→ = $PW of r = 1	Y2→ = $PW of r = 2	Y15→ = $PW of r = 15	Y...→ = $PW of r...	M1→ = average $PW of i = 1
240	0	12	...	101
0	112.5	312	...	45
180	375	202	...	222

300	300	324	...	311
150	225	223.5	...	214
0	112.5	456	...	202
unaccep. level	0	47	...	78
150	112.5	34.6	...	111
0	0	1023	...	444
120	300	233	...	211
180	75	231	...	311
0	75	65.8	...	56
unaccep. level	112.5	0	...	92
180	0	unaccep. level	...	22
270	37.5	122	...	111

Note that, n = 7 is an unacceptable attribute level for r = 1. The corresponding cell in the table above is therefore ignored in the calculation of the numerator and the denominator of M1,7. Next, calculate M2→.

2.1.3 Calculate σ

$\sigma = [(\sigma i)^2] = [(\sigma 1)^2 \quad (\sigma 2)^2]$, if I=2.

$(\sigma 1)^2$ = the square of the variance of segment i = 1

$\quad$ = Σ of the variance of n = 1 through n = N of segment 1

$\quad$ ={ $[1/(Si)] * [\Sigma r [Ind* (Y1,r - M1,1)^2]]$ $\quad$ +

$\quad$ $[1/(Si)] * [\Sigma r [Ind* (Y2,r - M1,2)^2]]$ $\quad$ + ... +

$\quad$ $[1/(Si)] * [\Sigma r [Ind* (YN,r - M1,N)^2]] \} * \{1/N\}$

Similarly, calculate $\sigma 2$.

37

2.1.4 Calculate F$\rightarrow$

$F_{i,r}$ is a value based upon $f_{i,n,r}$. There are $(R)*(I)$ $F_{i,r}$ values.

$F\rightarrow = (F1,1, F2,1, F1,2, F2,2,..., FI,R)$.

5    $(F1,1) = \Pi n\ [1/((2*pi*(\sigma 1)^{\wedge}2)^{\wedge}.5)]*[e^{\wedge}\ [-.5*(((Y1,1 - M1,1)^{\wedge}2/(\sigma 1)))]$

$= [1/((2*pi*(\sigma 1)^{\wedge}2)^{\wedge}.5)]*[e^{\wedge}\ [-.5*(((Y1,1 - M1,1)^{\wedge}2/(\sigma 1)))]*$

$[1/((2*pi*(\sigma 1)^{\wedge}2)^{\wedge}.5)]*[e^{\wedge}\ [-.5*(((Y2,1 - M1,2)^{\wedge}2/(\sigma 1)))]* ... *$

$[1/((2*pi*(\sigma 1)^{\wedge}2)^{\wedge}.5)]*[e^{\wedge}\ [-.5*(((Yn,1 - M1,n)^{\wedge}2/(\sigma 1)))]$

10   Similarly, calculate $(F2,1)$, $(F1,2)$, $(F2,2)$, $(F1,...)$, $(F2,...)$, $(F1,r)$, and $(F2,r)$.

2.1.5 Calculate posterior probability

Calculate $P1,1 = (S1)*(F1,1) / [\ \Sigma i\ (Si*Fi,1)]$. Also calculate $(P2,1)$, $(P1,2)$,

15    $(P2,2)$, $(P1,...)$, $(P2,...)$, $(P1,t)$, and $(P2,r)$.

2.1.6 Calculate the likelihood L

$L = \Sigma i,\ \Sigma r\ (\ ln(Fi,r) *\ Pi,r\ )$

20   ### 2.1.7 Iteration

Repeat steps 2.1.3 thru 2.1.6 in an iterative fashion until the posterior probabilities substantially converge.

2.1.7.1 Calculate new $\|M\|$

25   new $Mi,n = (\ \Sigma r\ Pi,r\ *\ Yn,r\ )/ [\ \Sigma r\ (Pi,r)]$

$Mi\rightarrow = (\ \Sigma r\ Pi,r\ *\ Yn,r\rightarrow\ )/ [\Sigma t\ (Pi,r)]$

As an example of this calculation, assume $P1,1 = .48$, $Y1,1 = 240$, and $Y1,2 = 0$. Accordingly, $M1,1 = (.48*240 + ...*0 + ...\ )/[\ .48 + ...] = 44$. Perform

30   similar calculation for all 15 rows of $M1\rightarrow$. Since $I = 2$, calculate $M2\rightarrow$ similarly.

2.1.7.2 Calculate new $(\sigma i)^2$

$(\sigma1)^2$ = the square of the variance of segment i = 1

= Σ of the square of the variance of n=1 through n=N of segment 1

= [1/(Σr P1,r)*N] * [Σr [(P1,r)*(Y1,r – M1,1)^2]] + [1/(Σr P1,r)*N]

* [Σr [(P1,r)*(Y2,r – M1,1)^2]] + ... + [1/(Σr P1,r)*N] * [Σr [(P1,r)

*(Yn,r – M1,1)^2]], where P is from 2.1.5 and M is from 2.1.7.1.

Since i = 2, also calculate $(\sigma2)^2$.

2.1.8 Ensure global Maximum

Repeat steps 2.1.1 through 2.1.7 using a new set of initial conditions.

2.1.9 Repeat for other numbers of segments I

Steps 2.1.1 thru 2.1.8 above were performed for I = 2. Repeat steps 2.1.1 thru 2.1.8 for I = 1, for I = 3, etc.

2.2 CAIC scoring

2.2.1 Calculate a CAIC score for I = 2 segments

Identify the final L (likelihood) of each of the two iterations. Identify the final L that is greater of the 2. Also identify the final posterior probabilities that correspond with that greater, final L.

CAIC = (-2)*(ln L) + (N)*(I)*(ln R)

2.2.2 Calculate a CAIC score for all other numbers of segments I

2.2.3 Identify the minimum CAIC score. The minimum score corresponds to the second optimal number of segments.

3. Determine Subgroups Based on First and Second Optimal Segments

Consider consumer r. A final posterior probability is defined for consumer r with respect to each segment of the first optimal number of segments. Again, a posterior probability reflects a probability that consumer r belongs to a particular segment. Similarly, a final posterior probability is defined for consumer r with respect to each segment of the second optimal number of segments.

Let $P^\$\to$ = the set of final posterior probabilities corresponding to r, with respect to each segment of the second optimal number of segments. $P^\$\to$ = $P^\$1,r = I, P^\$2,r = II,..., P^\$I,R = ...$

Let $P^u\to$ = the set of final posterior probabilities corresponding to r, with respect to each segment of the first optimal number of segments. $P^u\to$ = $P^u1,r = A, P^u2,r = B,..., P^uI,R = ...$

Multiply each final posterior probability of the set of $P^\$\to$ by each final posterior probability of the set of $P^u\to$. The combination of segments with the largest product of posterior probabilities is selected as a subgroup.

Segment	Probability		
I	0.7		
II	0.3		
A	0.2		
B	0.8		
	First Optimal # of segments		
Second Optimal # of segments		I	II
	A	0.14	0.06
	B	0.56	0.24

Accordingly, subgroup BI is identified as the optimal subgroup for consumer r =15. Repeat for all consumers r. If the largest product of posterior probabilities for two different consumers corresponds to a same set of $P^u i,r$ and

5 $P^\$i,r$, then the two consumers are placed into the same subgroup.

4. Determine Average Currency-normalized Preference Information for Each Subgroup

10

Calculate the $PW of each subgroup (= $U\Phi$), where Φ refers to the number of subgroups (AI, AII, BI, BII, ...). For example, assume that the first optimal number of segments is two (segment A and segment B), according to the CAIC score. Also, assume that the second optimal number of segments is two

15 (segment I and segment II). The final posterior probabilities are shown below for consumer r = 15.

To obtain a $PW matrix, $U\Phi\rightarrow$, for each of the subgroups, average the $PW for each consumer of the subgroup for a particular attribute level. Repeat

20 for every attribute level of every attribute ($U\Phi,n$).

5. Mix Preference Information

5.1 Determine optimal subgroup for current consumer

25 One of the advantages of latent cluster analysis is that a current consumer can be placed into one of the subgroups that were determined in step 3 without performing all the latent class iterations all over again, at least for the cases in which Si for all i are reasonably large. This condition ensures that the addition of the current consumer's preference information to the pool of past consumer

30 preference information doesn't really affect the subgroup definitions.

In other words, in order to place the current consumer in a subgroup, determine, for the current consumer, $P^\$cr\rightarrow$ for the first optimal number of segments AND determine $P^u cr\rightarrow$ for the second optimal number of segments, where cr = the current consumer.

5

5.1.1 Determine $P^u cr$

5.1.1.1 Calculate f→ for the current consumer

For the following calculation, use $\theta i\rightarrow$ of the final iteration that began with
10   the initial conditions which yielded the greatest final likelihood L, for the optimal number of segments I.

$$fi,n,cr = [(\theta i,n)^\wedge Yn,cr]*(1-\theta i,n)^\wedge(1-Yn,cr)$$
$$fi,cr\rightarrow = (fi,1,cr)$$
15   ...
$$(fi,n,cr).$$

5.1.1.2 Determine Fi,cr

Determine Fi,cr using the latent class analysis of step 1.

20

5.1.1.3 Calculate Posterior Probability Pcr→

Determine Pcr→ using the latent class analysis of step 2. In this regard, Pcr→ = $Pcr^u\rightarrow$. Accordingly, calculation of Pcr→ leads to calculation of $Pcr^u\rightarrow$.

25   ### 5.1.2 Determine $P^\$cr$

5.1.2.1 Calculate Fcr→

$$Fcr\rightarrow = (F1,cr, F2,cr, ..., FI,cr)$$
$$(F1,cr) = \Pi n\ [1/((2*pi*(\sigma1)^\wedge2)^\wedge.5)] * [e^\wedge [-.5 * (((Y1,cr - M1,1)^\wedge2/(\sigma1)^\wedge2))]$$
30        $$= [1/((2*pi*(\sigma1)^\wedge2)^\wedge.5)] * [e^\wedge [-.5 * (((Y1,cr - M1,1)^\wedge2/(\sigma1)^\wedge2))]*$$

42

$$[1/((2*pi*(\sigma 1)^2)^\wedge.5)] * [e^\wedge [-.5 * (((Y2,cr - M1,2)^\wedge 2/(\sigma 1)^2))]] * ... *$$

$$[1/((2*pi*(\sigma 1)^2)^\wedge.5)] * [e^\wedge [-.5 * (((YN,cr - M1,N)^\wedge 2/(\sigma 1)^2))]]$$

Since two segments have been deemed optimal, also calculate F2,cr.

5

5.1.2.2. Calculate posterior probability Pi,cr

P1,cr = (S1)*(F1,cr) / [Σi (Si*Fi,cr)]

Also calculate P2,cr because it has been deemed optimal to have two

10   segments. Thus, Pcr$$^\$$→ is calculated.

5.2 Determine optimal mixture of preference information associated with
consumer and average preference information associated with optimal subgroup

15   Because a one-to-one relationship exists between a correlation and a
regression, a correlation can be expressed as a regression. This can be
calculated with T→ = α*1→ + β*Pt→, where T→ = a column vector with 8
components = (8 trade-off question answers)*($PW conversion factor), 1→ =
8X1 column vector of 1's, Pt→ = 8X1 column vector (not to be confused with the

20   posterior probability vector) = (predicted answers based on the preference
information)* ($PW conversion factor), and α and β are scalar constants. Since
the resulting 8 equations include 2 unknowns, α and β, one can solve for α and
β.

         Similarly, an optimal mixture can be obtained using the equation T→ =

25   α*1→ + β (w*Pt→ + (1-w)Ps→), where Ps→ = 8X1 column vector = (predicted
answers based on the subgroup's average $PW information), w = a scalar that
represents the optimal mixture of current consumer and subgroup preference
information.

         As an example of the foregoing, assume that the actual preference

30   intensity (Pl→) and the predicted preference intensity (Ps→) shown below were

43

calculated during a validation process for the current consumer. Also assume that the $PW conversion factor = $50/util. Moreover, it should be assumed that subgroup BI is the optimal subgroup for the current consumer, and that Ps$\rightarrow$ for BI is calculated from the $PW of BI in the same manner as PWdiff$\rightarrow$ was

5   calculated from the preference information associated with the current consumer during validation.

T$\rightarrow$ = PI$\rightarrow$*conversion factor, where PI$\rightarrow$ is shown in table below.

Pt$\rightarrow$ = PWdiff$\rightarrow$*conversion factor, where Pt$\rightarrow$ is shown in table below.

10   I =[1 1 1 1 1 1 1 1]'

T[8,1] = { -150, 100, 50, -50, 50, -100, -50, 200 };

Pt[8,1] = { -61, 146, 61, -22, 10, -46, -90, 139 };

Ps[8,1] = { -176, 96.5, 45.5, 5.5, 61, -155.5, -61, 121.5 };

15

PI$\rightarrow$ (utils)	PWdiff$\rightarrow$ (utils)	T$\rightarrow$ ($)	Pt$\rightarrow$ ($)	Ps$\rightarrow$ for BI ($)
-3.00	-1.22	-150	-61	-176
2.00	2.92	100	146	96.5
1.00	1.22	50	61	45.5
-1.00	-0.44	-50	-22	5.5
1.00	0.2	50	10	61
-2.00	-0.92	-100	-46	-155.5
-1.00	-1.8	-50	-90	-61
4.00	2.78	200	139	-121.5

Using the above assumptions and data, w = 0.427. In other words, the optimal mixture is 42.7% current consumer preference information / 57.3% B1 subgroup preference information. It should be understood that the foregoing

20   value of w may be determined using the above information and any standard

44

statistical software program such as GAUSS 3.2, copyright 1995, 1996 Aptech Systems, Inc.

Recall that one subgroup was identified in step 3. Alternatively, more than one subgroup may be identified in step 3. For example, identified in step 3 may

5 be three subgroups corresponding to the highest three posterior probability products, subgroups associated with a highest $P^u cr \rightarrow$ (subgroups BI and BII in the above example), all subgroups corresponding to a posterior probability that is greater than a particular threshold value, or all determined subgroups. In the case that multiple subgroups are identified, $Ps \rightarrow$ may be calculated for each

10 subgroup and the foregoing equation may be used to calculate w for each subgroup. A subgroup having the best fit is then used to obtain the optimal mixture.

5.3 Compute final, stabilized currency-normalized preference information

15 associated with current consumer

For each attribute level which the current consumer has indicated as acceptable, compute a mixture of current consumer $PW and optimal subgroup $PW using w calculated in step 5.2. Specifically, the final, stabilized $PW matrix

20 for attribute levels that were acceptable to the current consumer = w*$PW matrix associated with current consumer + (1-w)*$PW matrix associated with optimal subgroup. Attribute levels that were unacceptable to the consumer remain unacceptable, regardless of the optimal subgroup.

25 **Detailed algorithm no. 2**

1.   Determine First Optimal Segments Based On (Un)Acceptable Attribute Levels

30

1.1 Definitions and Notation

N = the total number of attribute levels that are either acceptable or unacceptable to a respondent.

n = a specific attribute level, $n=1,...,N$

5 I = number of acceptable/unacceptable segments

i = a specific segment, $i=1,...,I$

T = total number of respondents for all segments

t = a specific respondent, $t=1,...,T$

j = iteration

10 $z_t = (z_{1,t},...,z_{N,t})$; vector with ones (acceptable level) and zero (unacceptable level) for all levels of all attributes 1 to N for respondent t

$\theta^j = (\theta_1^j,..., \theta_I^j)$; matrix with mean of the N attribute levels of the I segments

$\theta_i^j = (\theta_{i,1}^j,..., \theta_{i,N}^j)'$; vector with the mean of N attribute levels for segment I

(' indicates transposition)

15 $q_{i,t}^j$ = Probability that respondent t belongs to segment i

1.2 Latent Cluster Analysis

Do for $I =1,...,$ specified number (see section 3.1);

20 ### 1.2.1 Initial Conditions

$L^0 = 0$

$\theta^1 = (\theta_1^1,..., \theta_N^1)'$; the mean of N attributes

$d_t = \Sigma_n (z_{n,t} - \theta_n^1)$; the difference between respondent t scores and the mean of that attribute level

25 (t): specific respondent ordered by d_t, so each respondent has two indices: t and

(t)

$k = T/I$

$q_{i,t}^1 = 1$ if (t) $\in \{k*(i-1)+1, k*(i-1)+2, ...,k*i\}$

30 $= 0$ otherwise

Example for 1.2.1:

Suppose there are four respondents (T=4) and two segments (I=2); hence k = 2. The table below displays the data. First column gives t, the respondent number. The sum of the differences between score and mean is given in the second column. The third column the respondent number based on the difference. Finally, column four and five give the posterior probabilities. The latter are determined as follows: $q_{1,1}^1 = 0$ because (4) (which is (t)) $\notin \{1,...,2\}$, and $q_{2,1}^1 = 1$ because (4) $\in \{3,...,4\}$.

t	d_t	(t)	$q_{1,t}^1$	$q_{2,t}^1$
1	3	4	0	1
2	-2	1	1	0
3	-1	2	1	0
4	0	3	0	1

1.2.2

Do for j =1,..., converged

Do for i =1,..., I;

Do for n = 1,...,N;

$\theta_{i,n}^j = \Sigma_t z_{n,t} q_{i,t}^j / \Sigma_t q_{i,t}^j$; $\theta_{i,n}^j$ should be between 0 and 1.

$f_{i,t}^j = \Sigma_n q_{i,t}^j (z_{n,t} \ln(\theta_{i,n}^j) + (1 - z_{n,t}) \ln(1- \theta_{i,n}^j))$; where ln is the natural logarithm

$Q_i^j = (1/T) \Sigma_t q_{i,t}^j$

$L^j = \Sigma_t \Sigma_i f_{i,t}^j$

$q_{i,t}^{j+1} = Q_i^j \exp(f_{i,t}^j)/ \Sigma_i Q_i^j \exp(f_{i,t}^j)$

47

End 1.2.2 when $|L^j - L^{j-1}| < \varepsilon$, where ε is a small number (say 0.0001).

5

$CAIC_I = -2*L + N*I*ln(T+1)$; where L is the L for the last iteration j.

$E_I = 1 - \Sigma_i \Sigma_t (- q_{i,t} * ln(q_{i,t}) / (T * ln(I))$; E_I is the entropy metric for

segments I.

is for the last iteration j.

10

2.   Determine Second Optimal Segments Based On Currency-Normalized Part Worth Values

2.1 Definitions and Notation

15   N = the total number of attribute levels that are either acceptable or unacceptable

to a respondent.

n = a specific attribute level, n=1,...,N

S = number of $-PW segments

s = a specific segment, i=1,...,S

20   T = total number of respondents for all segments

t = a specific respondent, t=1,...,T

j = iteration

$y_{n,t} = (y_{1,t},...,y_{N,t})$, $PW value for attribute level n for respondent t

$\mu^j = (\mu_1^j,...,\mu_S^j)$; matrix with mean of the N attribute levels of the S segments

25   $\mu_s^j = (\mu_{s,1}^j,...,\mu_{s,N}^j)'$; vector with the mean of N attribute levels for segment s

$\sigma^{2,j} = (\sigma_1^{2,j},...,\sigma_I^{2,j})$; vector with variances of each segment (2 means square)

$p_{s,t}^j =$ Probability that respondent t belongs to segment s

$H^0 = 0$

30   2.2 Latent Cluster Analysis

Do for S =1,..., specified number (see section 3.1);

48

2.2.1 Initial Conditions

$\mu^0 = (\mu_1{}^0,\ldots, \mu_N{}^0)'$; the mean of N attributes

$d_t = \Sigma_n (y_{n,t} - \mu_n{}^0) * I(y_{n,t} > 0)$; the difference between respondent t scores and the

5   mean of that attribute level; $I(.)$ is an indicator function that is equal to one

if argument is true and zero otherwise.

(t) = specific respondent ordered by d_t ; so each respondent has two indices:

t and (t)

$k = T/S$

10    $p_{s,t}{}^1 = 1$ if $(t) \in \{k*(s-1)+1, k*(s-1)+2\ldots,k*s\}$

      = 0 otherwise

$H^0 = 0$

2.2.2

15   Do for j =1,…, converged

        Do for s =1,…, S;

            Do for n = 1,…,N;

20            $\mu_{s,n}{}^j = \Sigma_t (y_{n,t}\, p_{s,t}{}^j / \Sigma_t\, p_{s,t}{}^j)\, I(y_{n,t}$ is an acceptable level), where $I(.)=1$ if argument is true else $I(.)=0$.

            $(\sigma_s{}^j)^2 = \Sigma_n \Sigma_t\, p_{s,t}{}^j\, (y_{n,t} - \mu_{s,n}{}^j)^2\, I(y_{n,t}$ is an acceptable

25   level)$/ N\Sigma_t\, p_{s,t}{}^j$, where $I(.)=1$ if argument is true else $I(.)=0$.

            $g_{s,t}{}^j = \Sigma_n\, p_{s,t}{}^j\, (\,-(y_{n,t} - \mu_{s,n}{}^j)^2 / (2*\sigma_s{}^j)^2 - 0.5*\ln(2*pi*(\sigma_s{}^j)^2)\,)$; where ln is the natural logarithm and pi the number

30   3.14…

$$P_s^j = (1/T) \Sigma_t \, p_{s,t}^j$$

$$H^j = \Sigma_t \, \Sigma_s \, g_{s,t}^j$$

5

$$p_{s,t}^{j+1} = P_s^j \exp(g_{s,t}^j) / \Sigma_s \, P_s^j \exp(g_{s,t}^j)$$

End 2.2.2 when $|H^j - H^{j-1}| < \varepsilon$, where ε is a small number (say 0.0001).

$$CAIC_S = -2*L + N*I*\ln(T+1)$$

10

$$E_S = 1 - \Sigma_s \Sigma_t \, (- p_{s,t} * \ln(p_{s,t}) / (T * \ln(S)); \; E_S \text{ is the entropy metric for}$$
segments S.

3.   Creation Of Subgroups

15

There are now I*S possible subgroups. Respondent t is associated with subgroup (i,s) based upon maximum$_{(i,s)}$ of ($q_{i,t} * p_{s,t}$). The $-PW structure of a subgroup is the mean of the $-PW's of the respondents belonging to that subgroup. In calculating the mean, any levels of respondent t that are

20   unacceptable are ignored. The exact algorithm is described below.

Specified number for I (or S) = 2, 3 or 4 depending on the number of respondents. The table in section 3.8 gives the value of I (or S); e.g. if there are 200 respondents, then I (or S) <= 2. From now on we drop "(or S)", but

25   calculations have to be done for both segmentation solutions.

Determine entropy for each segment solution, $E_2, \ldots, E_I$, and the maximum of these $E_m = \max\{ E_2, \ldots, E_I \}$. Also, determine $E_2/E_m, \ldots, E_I/E_m$. Refer to section 1.2.4 (and 2.2.4).

30

Determine the CAIC for 1,...,I segment solutions. Refer to section 1.2.3 (and 2.2.3).

Select the segment solutions for which $E_i/E_m > 0.75$ and $E_i > 0.7$ ($i = 2, ...,$

5   I).

Choose the segment solution with the smallest CAIC for which criteria in section 3.4 hold.

10   When the segmentation solution for (un)acceptable and $-PW are determined, assign based on the multiplication of the posterior probabilities ($q_{i,t}$ * $p_{s,t}$), each respondent to a subgroup.

The subgroups that will be used are those that have at least the minimum

15   number of respondents per subgroup, as defined in the foregoing table.

Number of respondents	Maximum # of segments	Minimum number of respondents per subgroup
100	2	25
200	2	50
500	3	75
1000	4	75
1500	4	100

4.   Updating Of Subgroups

20   In one embodiment, sections 1 to 3 of this example may be repeated periodically to include new respondents in the subgroups. Below is an example detailing when the repetitions may occur. Section 4.1 applies to a system in its initial stages. Section 4.2 refers to a steady state system.

Initially, create subgroups with first 100, 200, 500, 1000 and 1500 respondents.

From then on, create subgroups of the last 1500 respondents; redo this when

5   300 new respondents are available. For example, on March 4 the subgroups may be updated to include respondent numbers 25001 to 26500. On March 25, there are 300 new respondents so the subgroups are updated to include respondent numbers 25301 to 26800

10   **5.   Current Respondent**

One embodiment for adding a current respondent to a subgroup and then mixing preference information of past respondents with preference information of the current respondent.

15

5.1 Assign to Segment

Use final $\theta_{i,n}$, and Q_i to determine posterior probability for (un)acceptables for respondent c, with data $z_{n,c}$, for each segment.

20   $$f_{i,c} = \Sigma_n(z_{n,c} \ln(\theta_{i,n}) +(1 - z_{n,c}) \ln(1- \theta_{i,n})); \text{ where ln is the natural logarithm}$$

$$q_{i,c} = Q_i \exp(f_{i,c})/ \Sigma_i Q_i \exp(f_{i,c})$$

Use final $\mu_{s,n}$, σ_s^2, and P_s to determine posterior probability for \$-PW for

25   respondent c, with data $y_{n,c}$, for each segment.

$$g_{s,c} = \Sigma_n \{[-(y_{n,c} - \mu_{s,n})^2 / (2^*\sigma_s^2) - 0.5^*\ln(2^*pi^*\sigma_s^2)]^*I(y_{n,c} \text{ is an acceptable}$$
$$\text{level})\}; \text{ where } pi = 3.14..., I(.) =1 \text{ if argument is true, else } I(.)=0.$$

30   $p_{s,c} = P_s g_{s,c} / \Sigma_s P_s g_{s,c}$

Determine $i^* = \{i \mid \max(q_{i,c}), i = 1,\ldots, I\}$; this defined the (un)acceptable segment this respondent belongs to.

Use all subgroups defined by i^* and $s = 1,\ldots, S$, for stabilization.

5

5.2 Stabilization

5.2.1 Definitions and Notation

T = Trade-off scores of current respondent * Conversion factor to $-PW

10 Pc = Predicted results for trade-offs from the $-PW structure for the eight trade-offs of

current respondent

$P_{i^*,s}$ = Predicted results for trade-offs of respondents c, from the $-PW structure for the eight trade-offs of the subgroup (i^*,s)

15 α, β, w : Scalar constants.

5.2.2 Algorithm

Based on the following linear regression equation:

20 Do for $s = 1,\ldots, S$

$$T = \alpha + \beta (w^*Pc + (1-w) P_{i^*,s})$$
$$= \alpha + \beta w^*(Pc - P_{i^*,s}) + \beta P_{i^*,s}$$
$$= \alpha + \gamma^*(Pc - P_{i^*,s})) + \beta P_{i^*,s}$$

25 estimates for α, γ, and β are obtained; w is then calculated as $w = \gamma/\beta$

If $w > 1$, w is set to 1; if $w < 0$, w is set to 0.

Determine the R^2 of this solution.

30

Choose the subgroup i^*, s^* that generates the highest R^2.

53

Stabilized preference structure is then defined by

$_PW (stabilized) = w ($-PW of current respondent) + (1-w) *($-PW of

5   subgroup i*, s*)